Using GNBD with Global File System

Configuration and Administration

4.7

Global_Network_Block_Device
ISBN: N/A
Publication date: July 2008

This book provides an overview on using Global Network Block Device (GNBD) with Red Hat GFS for Red Hat Enterprise Linux 4.7.

Using GNBD with Global File System: Configuration and Administration

Copyright © 2008 Red Hat, Inc.

Copyright © 2008 Red Hat, Inc. This material may only be distributed subject to the terms and conditions set forth in the Open Publication License, V1.0 or later with the restrictions noted below (the latest version of the OPL is presently available at *http://www.opencontent.org/openpub/*).

Distribution of substantively modified versions of this document is prohibited without the explicit permission of the copyright holder.

Distribution of the work or derivative of the work in any standard (paper) book form for commercial purposes is prohibited unless prior permission is obtained from the copyright holder.

Red Hat and the Red Hat "Shadow Man" logo are registered trademarks of Red Hat, Inc. in the United States and other countries.

All other trademarks referenced herein are the property of their respective owners.

The GPG fingerprint of the security@redhat.com key is:

CA 20 86 86 2B D6 9D FC 65 F6 EC C4 21 91 80 CD DB 42 A6 0E

1801 Varsity Drive
Raleigh, NC 27606-2072
USA
Phone: +1 919 754 3700
Phone: 888 733 4281
Fax: +1 919 754 3701
PO Box 13588
Research Triangle Park, NC 27709
USA

Introduction vii
1. About This Guide vii
2. Audience vii
3. Software Versions vii
4. Related Documentation vii
5. Document Conventions viii
6. Feedback x
1. Using GNBD with Red Hat GFS 1
2. Considerations for Using GNBD with Device-Mapper Multipath 3
1. Linux Page Caching 3
2. Fencing GNBD Server Nodes 3
3. GNBD Driver and Command Usage 5
1. Exporting a GNBD from a Server 5
2. Importing a GNBD on a Client 8
4. Running GFS on a GNBD Server Node 9
Index 11

Introduction

1. About This Guide

This book describes how to use Global Network Block Device (GNDB) with Global File System (GFS), including information about device-mapper multipath, GNDB driver and command usage, and running GFS on a GNBD server node.

2. Audience

This book is intended to be used by system administrators managing systems running the Linux operating system. It requires familiarity with Red Hat Enterprise Linux and GFS file system administration.

3. Software Versions

Software	Description
RHEL4	refers to RHEL4 and higher
GFS	refers to GFS 6.1 and higher

Table 1. Software Versions

4. Related Documentation

For more information about using Red Hat Enterprise Linux, refer to the following resources:

- *Red Hat Enterprise Linux Installation Guide* — Provides information regarding installation of Red Hat Enterprise Linux.
- *Red Hat Enterprise Linux Introduction to System Administration* — Provides introductory information for new Red Hat Enterprise Linux system administrators.
- *Red Hat Enterprise Linux System Administration Guide* — Provides more detailed information about configuring Red Hat Enterprise Linux to suit your particular needs as a user.
- *Red Hat Enterprise Linux Reference Guide* — Provides detailed information suited for more experienced users to reference when needed, as opposed to step-by-step instructions.
- *Red Hat Enterprise Linux Security Guide* — Details the planning and the tools involved in creating a secured computing environment for the data center, workplace, and home.

For more information about Red Hat Cluster Suite for Red Hat Enterprise Linux, refer to the following resources:

- *Red Hat Cluster Suite Overview* — Provides a high level overview of the Red Hat Cluster Suite.
- *Configuring and Managing a Red Hat Cluster* — Provides information about installing, configuring and managing Red Hat Cluster components.
- *Global File System: Configuration and Administration* — Provides information about installing, configuring, and maintaining Red Hat GFS (Red Hat Global File System).
- *LVM Administrator's Guide: Configuration and Administration* — Provides a description of the Logical Volume Manager (LVM), including information on running LVM in a clustered environment.
- *Using Device-Mapper Multipath* — Provides information about using the Device-Mapper Multipath feature of Red Hat Enterprise Linux.
- *Linux Virtual Server Administration* — Provides information on configuring high-performance systems and services with the Linux Virtual Server (LVS).
- *Red Hat Cluster Suite Release Notes* — Provides information about the current release of Red Hat Cluster Suite.

Red Hat Cluster Suite documentation and other Red Hat documents are available in HTML and PDF versions online at the following location:

http://www.redhat.com/docs

5. Document Conventions

Certain words in this manual are represented in different fonts, styles, and weights. This highlighting indicates that the word is part of a specific category. The categories include the following:

`Courier font`
Courier font represents `commands`, `file names and paths`, and `prompts` .

When shown as below, it indicates computer output:

```
Desktop       about.html       logs      paulwesterberg.png
Mail          backupfiles      mail      reports
```

`bold Courier font`
Bold Courier font represents text that you are to type, such as: **`service jonas start`**

If you have to run a command as root, the root prompt (#) precedes the command:

```
# gconftool-2
```

italic Courier font
: Italic Courier font represents a variable, such as an installation directory: *install_dir*/bin/

bold font
: Bold font represents **application programs** and **text found on a graphical interface**.

: When shown like this: **OK** , it indicates a button on a graphical application interface.

Additionally, the manual uses different strategies to draw your attention to pieces of information. In order of how critical the information is to you, these items are marked as follows:

Note

A note is typically information that you need to understand the behavior of the system.

Tip

A tip is typically an alternative way of performing a task.

Important

Important information is necessary, but possibly unexpected, such as a configuration change that will not persist after a reboot.

Caution

A caution indicates an act that would violate your support agreement, such as recompiling the kernel.

Warning

A warning indicates potential data loss, as may happen when tuning hardware for maximum performance.

6. Feedback

If you spot a typo, or if you have thought of a way to make this manual better, we would love to hear from you. Please submit a report in Bugzilla (*http://bugzilla.redhat.com/bugzilla/*) against the component `rh-cs`.

Be sure to mention the manual's identifier:

```
rh-gfs(EN)-4.7 (2008-07-24T15:10)
```

By mentioning this manual's identifier, we know exactly which version of the guide you have.

If you have a suggestion for improving the documentation, try to be as specific as possible. If you have found an error, please include the section number and some of the surrounding text so we can find it easily.

Using GNBD with Red Hat GFS

GNBD (Global Network Block Device) provides block-level storage access over an Ethernet LAN. GNBD components run as a client in a GFS node and as a server in a GNBD server node. A GNBD server node exports block-level storage from its local storage (either directly attached storage or SAN storage) to a GFS node.

Table 1.1, "GNBD Software Subsystem Components" summarizes the GNBD software subsystems components.

Software Subsystem	Components	Description
GNBD	`gnbd.ko`	Kernel module that implements the GNBD device driver on clients.
	`gnbd_export`	Command to create, export and manage GNBDs on a GNBD server.
	`gnbd_import`	Command to import and manage GNBDs on a GNBD client.
	`gnbd_serv`	A server daemon that allows a node to export local storage over the network.

Table 1.1. GNBD Software Subsystem Components

You can configure GNBD servers to work with device-mapper multipath. GNBD with device-mapper multipath allows you to configure multiple GNBD server nodes to provide redundant paths to the storage devices. The GNBD servers, in turn, present multiple storage paths to GFS nodes via redundant GNBDs. When using GNBD with device-mapper multipath, if a GNBD server node becomes unavailable, another GNBD server node can provide GFS nodes with access to storage devices.

This document how to use GNBD with Red Hat GFS and consists of the following chapters:

- *Chapter 2, Considerations for Using GNBD with Device-Mapper Multipath*
- *Chapter 3, GNBD Driver and Command Usage*
- *Chapter 4, Running GFS on a GNBD Server Node*

Considerations for Using GNBD with Device-Mapper Multipath

GNBD with device-mapper multipath allows you to configure multiple GNBD server nodes (nodes that export GNBDs to GFS nodes) to provide redundant paths to the storage devices. The GNBD server nodes, in turn, present multiple storage paths to GFS nodes via redundant GNBDs. When using GNBD with device-mapper multipath, if a GNBD server node becomes unavailable, another GNBD server node can provide GFS nodes with access to storage devices.

If you are using GNBD with device-mapper multipath, you need to take the following into consideration:

- Linux page caching, as desribed in *Section 1, "Linux Page Caching"*.
- Fencing GNBD server nodes, as described in *Section 2, "Fencing GNBD Server Nodes"*.
- GNBD device names; export names for GNBD devices must be unique. Additionally, you must specify the `-u` or `-U` when using the `gnbd_export` command. Exporting GNBD devices is described in *Chapter 3, GNBD Driver and Command Usage*.

1. Linux Page Caching

For GNBD with device-mapper multipath, do *not* specify Linux page caching (the `-c` option of the `gnbd_export` command). All GNBDs that are part of a logical volume must run with caching *disabled*. Data corruption occurs if the GNBDs are run with caching enabled. Refer to *Section 1, "Exporting a GNBD from a Server"* for more information about using the `gnbd_export` command for GNBD with device-mapper multipath.

2. Fencing GNBD Server Nodes

GNBD server nodes must be fenced using a fencing method that physically removes the nodes from the network. To physically remove a GNBD server node, you can use any fencing device: *except* the following: `fence_brocade` fence agent, `fence_vixel` fence agent, `fence_mcdata` fence agent, `fence_sanbox2` fence agent, `fence_scsi` fence agent. In addition, you *cannot* use the GNBD fencing device (`fence_gnbd` fence agent) to fence a GNBD server node. For information about configuring fencing for GNBD server nodes, refer to the *Global File System* manual.

GNBD Driver and Command Usage

The Global Network Block Device (GNBD) driver allows a node to export its local storage as a GNBD over a network so that other nodes on the network can share the storage. Client nodes importing the GNBD use it like any other block device. Importing a GNBD on multiple clients forms a shared storage configuration through which GFS can be used.

The GNBD driver is implemented through the following components.

- `gnbd_serv` — Implements the GNBD server. It is a user-space daemon that allows a node to export local storage over a network.
- `gnbd.ko` — Implements the GNBD device driver on GNBD clients (nodes using GNBD devices).

Two user commands are available to configure GNBD:

- `gnbd_export` (for servers) — User program for creating, exporting, and managing GNBDs on a GNBD server.
- `gnbd_import` (for clients) — User program for importing and managing GNBDs on a GNBD client.

1. Exporting a GNBD from a Server

The `gnbd_serv` daemon must be running on a node before it can export storage as a GNBD. You can start the `gnbd_serv` daemon running `gnbd_serv` as follows:

```
#gnbd_serv
gnbd_serv: startup succeeded
```

Once local storage has been identified to be exported, the `gnbd_export` command is used to export it.

Warning

When you configure GNBD servers with device-mapper multipath, you must not use page caching. All GNBDs that are part of a logical volume must run with caching disabled. By default, the `gnbd_export` command exports with caching turned off.

Note

A server should not import the GNBDs to use them as a client would. If a server exports the devices uncached, the underlying devices may also be used by `gfs`.

Usage

```
gnbd_export -dpathname-egnbdname [-c][-u][-U
```

`pathname`
Specifies a storage device to export.

`gnbdname`
Specifies an arbitrary name selected for the GNBD. It is used as the device name on GNBD clients. This name must be unique among all GNBDs exported in a network.

`-o`
Export the device as read-only.

`-c`
Enable caching. Reads from the exported GNBD and takes advantage of the Linux page cache.

By default, the `gnbd_export` command does *not* enable caching.

Warning

When you configure GNBD servers with device-mapper multipath, do not specify the `-c` option, as this lead sto data corruption. All GNBDs that are part of a logical volume must run with caching disabled.

Note

If you have been using GFS 5.2 or earlier and do *not* want to change your GNBD setup you *should* specify the `-c` option. Before GFS Release 5.2.1, Linux caching was enabled by default for `gnbd_export`. If the `-c` option is *not* specified, GNBD runs with a noticeable performance decrease. Also, if the `-c` option is *not* specified, the exported GNBD runs in timeout mode, using the default timeout value (the `-t` option). For more information about the `gnbd_export` command and its options, refer to the `gnbd_export` man page.

`-u`*`uid`*

Manually sets the Universal Identifier for an exported device. This option is used with `-e`. The UID is used by device-mapper multipath to determine which devices belong in a multipath map. A device must have a UID to be multipathed. However, for most SCSI devices the default Get UID command, `/usr/sbin/gnbd_get_uid`, will return an appropriate value.

Note

The UID refers to the device being exported, not the GNBD itself. The UIDs of two GNBD devices should be equal, only if they are exporting the same underlying device. This means that both GNBD servers are connected to the same physical device.

Warning

This option should only be used for exporting shared storage devices, when the `-U`*`command`* option does not work. This should almost never happen for SCSI devices. If two GNBD devices are not exporting the same underlying device, but are given the same UID, data corruption will occur.

`-U`*`Command`*

Gets the UID command. The UID command is a command the `gnbd_export` command will run to get a Universal Identifier for the exported device. The UID is necessary to use device-mapper multipath with GNBD. The command must use the full path of any executeable that you wish to run. A command can contain the %M, %m or %n escape sequences. %M will be expanded to the major number of the exported device, %m will be expaned to the minor number of the exported device, and %n will be expanded to the `sysfs` name for the device. If no command is given, GNBD will use the default command `/usr/sbin/gnbd_get_uid`. This command will work for most SCSI devices.

Examples

This example is for a GNBD server configured with GNBD multipath. It exports device `/dev/sdc2` as GNBD `gamma`. Cache is disabled by default.

```
gnbd_export -d /dev/sdc2 -e gamma -U
```

This example is for a GNBD server *not* configured with GNBD multipath. It exports device `/dev/sdb2` as GNBD `delta` with cache enabled.

```
gnbd_export -d /dev/sdb1 -e delta -c
```

This example exports device `/dev/sdb2` as GNBD `delta` with cache enabled.

```
gnbd_export -d /dev/sdb2 -e delta -c
```

2. Importing a GNBD on a Client

The `gnbd.ko` kernel module must be loaded on a node before it can import GNBDs. When GNBDs are imported, device nodes are created for them in `/dev/gnbd/` with the name assigned when they were exported.

Usage

```
gnbd_import -iServer
```

Server
: Specifies a GNBD server by hostname or IP address from which to import GNBDs. All GNBDs exported from the server are imported on the client running this command.

Example

This example imports all GNBDs from the server named `nodeA`.

```
gnbd_import -i nodeA
```

Running GFS on a GNBD Server Node

You can run GFS on a GNBD server node, with some restrictions. In addition, running GFS on a GNBD server node reduces performance. The following restrictions apply when running GFS on a GNBD server node.

Important

When running GFS on a GNBD server node you *must* follow the restrictions listed; otherwise, the GNBD server node will fail.

1. A GNBD server node must have local access to all storage devices needed to mount a GFS file system. The GNBD server node must not import (`gnbd_import` command) other GNBD devices to run the file system.
2. The GNBD server must export all the GNBDs in uncached mode, and it must export the raw devices, not logical volume devices.
3. GFS must be run on top of a logical volume device, not raw devices.

Note

You may need to increase the timeout period on the exported GNBDs to accommodate reduced performance. The need to increase the timeout period depends on the quality of the hardware.

Index

D

device-mapper multipath, 3
    fencing GNBD server nodes, 3
    Linux page caching, 3
driver and command usage, 5
    exporting from a server, 5
    importing on a client, 8

E

exporting from a server daemon, 5

F

feedback, x, x
fencing GNBD server nodes, 3

G

GFS, using on a GNBD server node, 9
GNBD, using with Red Hat GFS, 1
gnbd.ko module, 5, 8
gnbd_export command , 5, 6
gnbd_import command , 5, 8
gnbd_serv daemon, 5, 5

I

importing on a client module, 8

L

Linux page caching, 3

S

software subsystem components, 1